SPR Day 16

Agenda:

1: Q&A about new "online" flipped classroom format
2: Short exercises about Gaussian mixture models and EM
3: Q&A like office hours

Rules
1) Mute except to talk (space bar to unmute)

will release slides

Welcome

Hi!

videos released ahead

update HW3 CP3 on schedule

Questions?

HW, CP as planned
quizzes 15 min w/ same content
final exam (like midterm)

Suggestions:

- Questions abt specific time in video
- Little chunks good ✓
- Live blackboard
- Provide more landmarks in notes
- Thinner ink
- Less zoom in/out

Exercise 1: Write two versions of GMM likelihood, complete and incomplete

single data point

$x_n \in \mathbb{R}^D$

incomplete

$\log p(x_n \mid \pi, \mu, \sigma)$

$\sigma = \{\sigma_{kd}\}$ $k = 1 \ldots K$ $d = 1 \ldots D$

variance separate/indep ~~for~~ each component, dim

complete

$\log p(x_n, z_n \mid \pi, \mu, \sigma)$

incomplete

$\sum_k$ cover $D \times D$ for each k

$$\log p(x_n \mid \pi, \mu, \tau) = \log \sum_{k=1}^{K} \pi_k \underbrace{\prod_{d=1}^{D} \text{NormPDF}(x_{nd} \mid \mu_{kd}, \tau_{kd}^2)}_{\text{MVNormPDF}(x_n \mid \mu_k, \text{diag}(\tau_k^2))}$$

univariate

$$p(x_n) = \text{GMMPDF}(x_n)$$

$$\text{MVNormPDF}(x_n \mid \mu_k, \Sigma_k)$$

complete

$$\log p(x_n, z_n \mid \pi, \mu, \tau) = \sum_k z_{nk} \log \pi_k + \sum_{k=1}^{K} z_{nk} \underbrace{\sum_{d=1}^{D} \log \text{NormPDF}(x_n \mid \mu_{kd}, \tau_{kd}^2)}$$

$$z_n \sim \text{Cat}(\pi_1, \dots, \pi_K)$$

$z_n = [z_{n1} \dots z_{nK}]$

1 0000

00100

$$p(x_n \mid z_n) = \prod_{k=1}^{K} \text{NormPDF}(x_n \mid \mu_k, \tau_k)^{\boxed{z_{nk}}}$$

one hot

$z_{n3} = 1$

$= 1 \cdot 1 \cdot \text{NormPDF}(\cdot \mid \mu_3, \tau_3) \cdot 1 \cdot 2$

Exercise 2: What is entropy?

- Write down the expectation that defines the entropy of a discrete random variable z with PMF q(z)

$q(z=1)$
$\vdots$
$q(z=K)$

Sample space of z = {1, 2, 3, … K}

$\mathrm{Entropy}[q(z)] =$ ______________

$$\text{Entropy}[q(z)] = -\sum_{z \in \Omega = \{1,\dots,K\}} q(z) \log q(z)$$

$$= \mathbb{E}_{q(z)}[-\log q(z)]$$

$q(z) = \text{Cat}(0.99, 0.01, 0.01)$

low entropy

optional readings on schedule

PRML Section 1.6

high entropy = lots of uncertainty

low entropy : certain outcome

Exercise 3: What is KL divergence?

- Write down the expectation that defines the KL divergence between two distributions over a discrete random variable: q(z) and p(z)

$$\overrightarrow{q \qquad\qquad p}$$

$$KL(q(z)\,||\,p(z)) = -\sum_{\substack{\cancel{z\in\Omega} \\ z\in\{1,2,\dots K\}}} q(z)\log\frac{p(z)}{q(z)}$$

from q to p

Bishop 1.113

$$= \mathbb{E}_q\left[-\log\frac{p(z)}{q(z)}\right] = \mathbb{E}_q\left[\log\frac{q(z)}{p(z)}\right]$$

1) KL Purpose

how ~~close~~ close is q to p

smaller value : close

larger → far

0.0

5.8

17.23

2) $KL \geq 0$ always positive

3) $KL = 0$ iff $q = p$

$KL(q\|p) \neq KL(p\|q)$

Exercise 4

Assume:
- GMM for 1-dim data, with K = 3
- There exists a function scipy.stats.norm.logpdf that computes the log pdf of a 1d normal

Given:
- x : a scalar observation from a GMM
- pi_K, mu_K, sigma_K : GMM parameters

Goal:

Write python function that could compute

log p(x) = log GMMPDF(x)

EXTRA POINTS: Should be numerically stable!

Exercise 4: Solution

```
def calc_log_lik(x, pi_K, mu_K, sig_K):
                         vec K
    return np.log(np.sum(pi_K *
        scipy.stats.norm.pdf(x, mu_K, sig_K)))
                  vec K
```

logsumexp trick!

$$\log\Big(\sum_k \pi_k \,\text{NormPDF}(x \mid \mu_k, \sigma_k)\Big)$$

if x far from μ_k

small

$$\boxed{\log \sum_k e^{a_k}}$$

vec K

logsumexp(np.log(pi_K) + norm.logpdf(x, mu_K, sig_K))

numerically stable, useful for CP3

why not just do Key Idea #2?

Proposal $\max_{\pi,\mu,\sigma^2} \mathbb{E}_q[\log p(x,z)]$

r defines $q(z)$

optimal $q(z_n|r_n) \neq p(z_n|x_n)$

① relation to incomplete?

② what would r look like?

learned $q(z|r)$ poor, needed!

What we do: $\max_{\pi,\mu,\sigma^2,r} \underbrace{\mathbb{E}_q[\log p(x,z)] - \mathbb{E}_q[\log q(z)]}_{\mathcal{L}(x,r,\pi,\mu,\sigma)}$

optimal $r_n \ni p(z_n|x_n) = q(z_n|r_n)$